

1649

PATENT  
1379-1-013IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS : María Pilar PRIETO-DAPENA *et al* #7  
SERIAL NO. : 09/600,848  
FILED : July 24, 2000  
FOR : PROMOTER AND REGULATOR SEQUENCES ha ds10 GI:  
A GENE LEA OF SUNFLOWERS EXPRESSED  
EXCLUSIVELY IN SEEDS FROM THE MATURATION  
PHASE

PETITION FOR GRANT OF PRIORITY UNDER 35 USC 119ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

RECEIVED

JAN 04 2001

TECH CENTER 1600/2800

Dear Sir:

Applicant hereby petitions for grant of priority of the present Application on the basis of the following prior filed foreign Application:

COUNTRY  
SPAINSERIAL NO.  
P 9800122FILING DATE  
JANUARY 23, 1998

To perfect Applicant's claim to priority, a certified copy of the above listed prior filed Application is enclosed.

Acknowledgment of Applicant's perfection of claim to priority is accordingly requested.

Respectfully submitted,

---

David A. Jackson  
Attorney for Applicant  
Registration No. 26,742KLAUBER & JACKSON  
411 Hackensack Avenue  
Hackensack, NJ 07601  
(201)487-5800



# 2

**OFICINA ESPAÑOLA**

de

**PATENTES y MARCAS****RECEIVED**

JAN 04 2001

TECH CENTER 1600/2900

# CERTIFICADO OFICIAL

Por la presente certifico que los documentos adjuntos son copia exacta de los que obran en el expediente de PATENTE de INVENCION número 9800122, de acuerdo con la concesión efectuada con fecha 2 de Agosto de 2000.

Madrid, 6 de octubre de 2000

El Director del Departamento de Patentes  
e Información Tecnológica.

P.D.

M. MADRUGA



OFICINA ESPAÑOLA DE PATENTES Y  
MARCAS

INSTANCIA DE SOLICITUD DE:

☒ PATENTE DE INVENCION ☐ MODELO DE UTILIDAD

(1) <input type="checkbox"/> SOLICITUD DE ADICION <input type="checkbox"/> SOLICITUD DIVISIONAL <input type="checkbox"/> CAMBIO DE MODALIDAD <input type="checkbox"/> TRANSFORMACION SOLICITUD EUROPEA		(2) EXPED. PRINCIPAL O DE ORIGEN MODALIDAD ..... NUMERO SOLICITUD ..... FECHA SOLICITUD ...../...../..... MODALIDAD ..... NUMERO SOLICITUD ..... FECHA SOLICITUD ...../...../.....		NUMERO DE SOLICITUD  FECHA Y HORA DE PRESENTACION EN O.E.P.M.  98 ENE 23 12:06  FECHA Y HORA DE PRESENTACION EN LUGAR DISTINTO O.E.P.M.	
(3) LUGAR DE PRESENTACION España		CODIGO ES			
(4) SOLICITANTE(S)		APELLIDOS O DENOMINACION JURIDICA		NOMBRE	
		Consejo Superior Investigaciones Científicas		Q28/18002D	
(5) DATOS DEL PRIMER SOLICITANTE					
DOMICILIO ..... Serrano, 117					
LOCALIDAD ..... MADRID					
PROVINCIA ..... MADRID					
PAIS RESIDENCIA ..... España					
NACIONALIDAD ..... Española					
TELEFONO ..... 585 50 00					
CODIGO POSTAL ..... 28006					
CODIGO PAIS ..... ES					
CODIGO NACION ..... ES					
(6) INVENTOR(ES)		(7) <input type="checkbox"/> EL SOLICITANTE ES EL INVENTOR <input checked="" type="checkbox"/> EL SOLICITANTE NO ES EL INVENTOR O UNICO INVENTOR		(8) MODO DE OBTENCION DEL DERECHO <input checked="" type="checkbox"/> INVENC. LABORAL <input type="checkbox"/> CONTRATO <input type="checkbox"/> SUCESION	
APELLIDOS		NOMBRE		NACIONALIDAD	
Prieto-Dapena		Pilar		ES	
Almoguera		Concepción		ES	
Jordano		Juan		ES	
(9) TITULO DE LA INVENCION					
Promotor y secuencias reguladoras de Ha ds10 G1 : un gen LEA de girasol expresado exclusivamente en semillas desde la fase de maduración.					
(10) INVENCION REFERENTE A PROCEDIMIENTO MICROBIOLOGICO SEGUN ART. 25.2 L.P. <input checked="" type="checkbox"/> SI <input type="checkbox"/> NO					
(11) EXPOSICIONES OFICIALES					
LUGAR ..... FECHA RECEIVED					
(12) DECLARACIONES DE PRIORIDAD					
PAIS DE ORIGEN		COD. PAIS	NUMERO	FECHA	
				JAN 04 2001	
				TECH CENTER 1600/2900	
(13) EL SOLICITANTE SE ACOGE A LA EXENCION DE PAGO DE TASAS PREVISTA EN EL ART. 162 L.P. <input checked="" type="checkbox"/> SI <input type="checkbox"/> NO					
(14) REPRESENTANTE		APELLIDOS		NOMBRE	
				CODIGO	
DOMICILIO		LOCALIDAD		PROVINCIA	
				COD. POSTAL	
(15) RELACION DE DOCUMENTOS QUE SE ACOMPAÑAN					
<input checked="" type="checkbox"/> DESCRIPCION. N.º DE PAGINAS...		<input checked="" type="checkbox"/> DOCUMENTO DE REPRESENTACION			
<input checked="" type="checkbox"/> REIVINDICACIONES. N.º DE PAGINAS...		<input type="checkbox"/> PRUEBAS			
<input checked="" type="checkbox"/> DIBUJOS. N.º DE PAGINAS...		<input checked="" type="checkbox"/> JUSTIFICANTE DEL PAGO DE TASAS			
<input checked="" type="checkbox"/> RESUMEN		<input type="checkbox"/> HOJA DE INFORMACIONES			
<input type="checkbox"/> DOCUMENTO DE PRIORIDAD		<input type="checkbox"/> COMPLEMENTARIAS			
<input type="checkbox"/> TRADUCCION DEL DOCUMENTO DE PRIORIDAD		<input type="checkbox"/> OTROS			
(16) NOTIFICACION DE PAGO DE LA TASA DE CONCESION					
Se le notifica que esta solicitud se considerará retirada si no procede al pago de la tasa de concesión: para el pago de esta tasa dispone de tres meses a contar desde la publicación del anuncio de la concesión en el BOPF, más los diez días que establece el art. 81 del R.D. 10-10-86.					
FIRMA DEL FUNCIONARIO				FIRMA DEL SOLICITANTE O REPRESENTANTE	
				Pedro Ojeda	

ILMO. SR. DIRECTOR DE LA OFICINA ESPAÑOLA DE PATENTES Y MARCAS

CUMPLIMENTAR LOS TRES EJEMPLARES SALVO ZONAS EN ROJO



# PATENTE

## RESUMEN Y GRAFICO

NUMERO DE SOLICITUD

2000122

FECHA DE PRESENTACION

'98 ENE 23 12:06

RESUMEN (Máx. 150 palabras)

**Promotor y secuencias reguladoras de Ha ds10 G1 : un gen LEA de girasol expresado exclusivamente en semillas desde la fase de maduración.**

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA-I de girasol, Ha ds10 G1. Estas secuencias presentan unas características muy apropiadas para su uso en la modificación de semillas (por ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante ejemplos como estudios de la acumulación y localización del ARNm Ha ds10 en el sistema homólogo. Estos estudios muestran tanto los elevados niveles de expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como sus absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol.

GRAFICO



31) NUMERO	DATOS DE PRIORIDAD		32) FECHA	33) PAIS	A1	12) PATENTE DE INVENCION
						21) NUMERO DE SOLICITUD
						22) FECHA DE PRESENTACION 23-01-98

71) SOLICITANTE(S) Consejo Superior Investigaciones Científicas			NACIONALIDAD Española	
DOMICILIO Serrano, 117 28006 MADRID				
72) INVENTOR(ES) Pilar Prieto-Dapena, Concepción Almoguera y Juan Jordano.				
73) TITULAR(ES) Consejo Superior Investigaciones Científicas				
11) N.º DE PUBLICACION	45) FECHA DE PUBLICACION	62) PATENTE DE LA QUE ES DIVISIONARIA	GRAFICO (SOLO PARA INTERPRETAR RESUMEN)	
51) Int. Cl. <b>F</b> C12N 15/82, 15/29, 2014 J/00			<b>RECEIVED</b>  JAN 04 2001  TECH CENTER 1600/2900	
54) TITULO  <b>Promotor y secuencias reguladoras de Ha ds10 G1 : un gen LEA de girasol expresado exclusivamente en semillas desde la fase de maduración.</b>				

## 57) RESUMEN (APORTACIÓN VOLUNTARIA, SIN VALOR JURÍDICO)

**Promotor y secuencias reguladoras de Ha ds10 G1 : un gen LEA de girasol expresado exclusivamente en semillas desde la fase de maduración.**

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA-I de girasol, Ha ds10 G1. Estas secuencias presentan unas características muy apropiadas para su uso en la modificación de semillas (por ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante ejemplos como estudios de la acumulación y localización del ARNm Ha ds10 en el sistema homólogo. Estos estudios muestran tanto los elevados niveles de expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como sus absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol.

## TÍTULO

**PROMOTOR Y SECUENCIAS REGULADORAS DE *HA DS10 G1*: UN GEN LEA DE GIRASOL EXPRESADO EXCLUSIVAMENTE EN SEMILLAS DESDE LA FASE DE MADURACIÓN.**

5

## SECTOR DE LA TÉCNICA

Agricultura. Esta invención se relaciona con la obtención de secuencias de ADN reguladoras ("promotores") y la construcción, usando dichas secuencias, de nuevos genes quiméricos capaces de expresarse de forma específica en semillas de plantas transgénicas. El gen *Ha ds10 G1* tiene la peculiaridad de expresarse exclusivamente en semillas de girasol desde la fase de maduración hasta la de desecación; sin responder a hormonas como el ácido abscísico (ABA), o al estrés hídrico en tejidos vegetativos. Además, el gen *Ha ds10 G1* se expresa de forma homogénea en embriones inmaduros, y preferentemente en el parénquima en empalizada de los cotiledones de embriones maduros. Estos patrones de expresión, junto con los elevados niveles de actividad del gen, sugieren que sus secuencias reguladoras sean especialmente adecuadas para la manipulación genética de sustancias de reserva en semillas

20

## ESTADO DE LA TÉCNICA

Para conferir expresión específica en semillas de plantas transgénicas, hasta el momento se han aislado, caracterizado y utilizado promotores pertenecientes sobre todo a genes vegetales que codifican proteínas de reserva, u otros productos expresados exclusivamente en semillas durante diversas etapas del desarrollo [véanse por ejemplo las siguientes referencias bibliográficas y patentes, así como otros documentos citados en ellas: Thomas TL, en *Plant Cell*, vol 5, pp 1401-1410, 1993; Gatehouse JA, y Shirsat AH, en *Control of Plant Gene Expression*, pp 357-375, CRC press, 1993; y las patentes USA números: 5530192, 5530194 y 5420034]. Esto ha permitido por ejemplo la obtención de nuevas plantas transgénicas con semillas modificadas en su contenido de ácidos grasos y de proteínas de reserva [véanse por ejemplo:

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE y Davies HM, en *Science*, vol. 257, pp.72-74, 1992; y Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, y Muntz K, en *Molecular and General Genetics* 242: 226-236, 1994]. Para el desarrollo del enorme potencial de esta

5 técnica, pudieran ser útiles otros promotores con distintas especificidades de tejido en la semilla y diversos patrones temporales de expresión. Recientemente en nuestro grupo, y otros laboratorios, hemos descrito la expresión en semillas de genes que codifican proteínas de choque térmico de bajo peso molecular (sHSPs: *small heat-shock proteins*). Uno de estos genes, *Ha hsp17.7 G4*,

10 muestra, en plantas transgénicas de tabaco, patrones de expresión adecuados para su posible uso en la modificación de semillas mediante ingeniería genética: dicho gen se expresa desde etapas tempranas de la maduración de la semilla, y con una especificidad de tejido asociada a los cotiledones [Coca MA, Almoguera C, Thomas TL, y Jordano J, en: *Plant Molecular Biology* 31: 863-876, 1996]. Sin

15 embargo el gen *Ha hsp17.7 G4*, al igual que otros genes vegetales sHSP expresados en semillas, también se expresa en respuesta al calor (choques térmicos) en tejidos vegetativos de la planta tras la germinación de la semilla. Esto último imposibilita su uso en ingeniería genética cuando se requieren

20 secuencias de ADN reguladoras que garanticen que no haya expresión de los genes quiméricos fuera de la semilla: por ejemplo, cuando la expresión fuera de lugar de estos genes pueda afectar a la viabilidad, el crecimiento o la salubridad de las plantas transgénicas. Para solucionar estos problemas hemos modificado

25 las secuencias reguladoras del gen *Ha hsp17.7 G4* de forma que genes quiméricos que contengan estas secuencias mantengan su expresión en semillas y pierdan su inducción por calor; procedimiento utilizable para la modificación y uso similar de secuencias reguladoras de otros genes sHSP expresados en semillas [Almoguera, Prieto-Dapena y Jordano, solicitud de

30 patente #9602746 (Oficina Española de Patentes)]. De forma alternativa, también hemos propuesto un uso similar para el promotor y las secuencias reguladoras del gen de girasol *Ha hsp17.6 G1*, que únicamente se expresa en semillas. Dicho gen no responde al calor o a otro tipo de estrés (frío, desecación, tratamiento hormonal con ABA) en tejidos vegetativos [Carranco,

Almoguera y Jordano, solicitud de patente #9701215 (Oficina Española de Patentes).

En la presente solicitud proponemos usos análogos alternativos para el promotor y las secuencias reguladoras del gen LEA de girasol *Ha ds10 G1*. El gen *Ha ds10 G1* está incluido en un clon genómico correspondiente a un ADNc descrito previamente (*Ha ds10*, número de acceso X50699) cuyos patrones de expresión se conocían de forma incompleta [Almoguera y Jordano, *Plant Mol. Biol.* 19:781-792, 1992]. El promotor y secuencias reguladoras de este gen (*Ha ds10 G1*) han sido clonados y se describen, caracterizan y utilizan por primera vez en los ejemplos de esta solicitud. El gen *Ha ds10 G1* pertenece a la familia de genes LEA (*Late Embryogenesis Abundant*) de Clase I (tipo D-19 ó LEA-I). Estos genes codifican proteínas altamente conservadas en varias especies vegetales, y su expresión está generalmente restringida a semillas y a fases tempranas de la germinación [ver por ejemplo las siguientes revisiones: Dure III, L., *Structural motifs in Lea proteins*, en *Plant Responses to Plant Dehydration During Environmental Stress*, Close TJ and Bray EA Eds., *Current Topics in Plant Physiology* 10: 91-103, 1993; y Delseny M, Gaubier P, Hull G, Saez-Vasquez J, Gallois P, Raynal M, Cooke R, Grellet F., *Nuclear Genes expressed during seed desiccation: relationship with responses to stress*, en *Stress-induced Gene Expression in Plants* (Batra, A. S., ed.), pp. 25-59, Harwood Academic Publishers, Reading, 1994]. Los promotores de los genes LEA no han sido considerados como buenos candidatos para su uso en proyectos de modificación de sustancias de reserva en semilla, ya que en general presentan actividad en fases posteriores a la maduración de la semilla, como durante la desecación del embrión [ver las consideraciones de Kridl JC, Knauf VC, Thompson GA, en *Control of Plant Gene Expression*. pp. 481-498, CRC press, 1993]. Sin embargo se conocen genes LEA que se activan en fases de maduración anteriores a la desecación, como los genes de algodón denominados LEA-A [Hughes DW y Galau GA, *The Plant Cell* 3:605-618, 1991]. También dentro los genes LEA de clase I se conocen ejemplos de activación anterior a la desecación, como en el caso de los genes *At Em1*, *emb564*, y *emb1* [respectivamente en arábido, maíz y zanahoria: Gaubier P, Raynal M,

Hull G, Huestis GM, Grellet F, Arenas C, Pages M, y Delseny M, *Mol. Gen. Genet.*, 238: 409-418, 1993; Williams B, y Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991; Wurtele ES, Wang H, Durgerian S, Nikolau BJ, y Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. Estos ejemplos indicarían el posible uso de

5 secuencias reguladoras de genes de esta familia para la modificación de semillas. No obstante, su uso concreto estaría limitado tanto por los niveles de expresión alcanzados en cada caso y en cada fase del desarrollo; como por las distintas especificidades de tejido. Así por ejemplo, aunque en *Arabidopsis* el gen *At Em1* se activa tempranamente, su expresión esta fundamentalmente

10 restringida a tejidos provasculares de los cotiledones y a tejidos corticales externos del eje embrionario [Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., y Delseny, M., *Mol. Gen. Genet.*, 238: 409-418, 1993]. En el caso del gen *emb1* de zanahoria, sus ARNm se localizan preferente en los meristemos del embrión, particularmente en el *procambium*

15 [Wurtele ES, Wang H, Durgerian S, Nikolau BJ, y Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. No se han publicado las secuencias genómicas del gen *emb564*, y se desconoce la localización precisa de sus ARNm [Williams B, y Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991].

La expresión del gen de girasol *Ha ds10 G1*, así como su promotor y

20 secuencias reguladoras presentan, como se describe a continuación, unas características únicas entre las de otros miembros de la familia LEA-I; lo que hace que dichas secuencias sean potencialmente utilizables en la modificación de semillas mediante ingeniería genética.

## 25 DESCRIPCION DE LA INVENCION

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA-I de girasol, *Ha ds10 G1*. Estas secuencias (Ejemplo 1) presentan unas características muy apropiadas para su uso en la modificación de semillas (por

30 ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante otros ejemplos: A.- Estudios de la acumulación y localización del ARNm *Ha ds10* en el sistema homólogo (Ejemplo

2). Estos estudios muestran tanto los elevados niveles de expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como su absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol. B.- En el ejemplo 3, ilustramos también el posible uso de dichas secuencias mediante la construcción y análisis en plantas transgénicas de distintos genes quiméricos; usando el promotor y combinaciones de distintas secuencias reguladoras de *Ha ds10 G1* (5'-flanqueantes, codificantes, intrón y 3'-flanqueantes), con el gen indicador (*reporter*) de la  $\beta$ -glucuronidasa bacteriana (GUS). Estos ejemplos demuestran en un sistema heterólogo modelo (tabaco) la utilidad de los distintos genes quiméricos ensayados: alto nivel de expresión y especificidad de semillas desde fases tempranas de la maduración, así como la contribución funcional de las distintas secuencias ensayadas. Mediante los ejemplos adjuntos mostramos que la especificidad de semillas está conferida fundamentalmente por el promotor y secuencias 5'-flanqueantes de *Ha ds10G1* (incluyendo secuencias no-transcritas y transcritas: como el 5'-UTR y parte de la secuencia codificante). Adicionalmente las secuencias 3'-flanqueantes incrementan los niveles de expresión en semillas; y el intrón los reduce de forma específica en tejidos no-embrionarios. Dada la conservación de la regulación de la expresión de genes embrionarios en semillas de plantas, incluidos los genes LEA-I [Thomas TL, en *The Plant Cell* 5:1401-1410, 1993]; estas secuencias podrían usarse tanto en el sistema homólogo (el girasol) como en otros sistemas heterólogos de gran importancia económica (por ejemplo la colza, la soja, el maíz, etc).

La realización práctica de esta invención, representada con los ejemplos y figuras adjuntos, utiliza técnicas convencionales de Biología Molecular, Microbiología, ADN recombinante; y de producción de plantas transgénicas, que son de uso común en laboratorios especializados en estos campos. Estas técnicas están explicadas con suficiente detalle en la literatura científica [veáanse por ejemplo: Sambrook J, Fritsch EF, y Maniatis T, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor laboratory Press, 2ª Edición, 1989;

Glover DM, *DNA Cloning*, IRL Press , 1985; Lindsey K., *Plant Tissue Culture Manual*, Kluwer Academic Publishers, 1993; y Gelvin SB, Schilperoort RA, Verma DPS, *Plant Molecular Biology Manual*, Kluwer Academic Publishers , 1992]. Para otros detalles mas específicos, se citan las referencias bibliográficas

5 pertinentes en el lugar correspondiente de esta solicitud.

EJEMPLO 1: clonación, determinación del mapa de restricción, secuencia nucleotídica, y análisis del promotor de *Ha ds10 G1*.

Para obtener el clon *Ha ds10 G1* se rastreó la genoteca de ADN genómico de girasol descrita por Coca *et al.* [Plant Mol. Biol. 31: 863-876, 1996], con la sonda correspondiente al ADNc completo *Ha ds10* [Almoguera y Jordano, Plant Mol. Biol. 19: 781-792, 1992]; usando las condiciones de hibridación y procedimientos estandard de clonación molecular descritos con suficiente detalle en la primera de estas referencias (Coca *et al.*, 1996). Así, aislamos un fago (IGEM11) con un inserto de ADN genómico de girasol de aproximadamente 16.5 Kb cuyo mapa parcial se muestra en la Figura 1. Mediante análisis de restricción, determinamos que dos fragmentos adyacentes de Sac I (de 4.2 y 9.3 Kb) contienen las secuencias que hibridan con el ADNc. Se determinó un mapa de restricción detallado del primero de estos fragmentos y de parte (4 Kb) del segundo (Figura 1). Distintos subfragmentos de ADN genómico, correspondientes a la región mapeada, se clonaron en el vector pBluescript SK+, dando lugar a los plásmidos cuyo nombre e inserto se indica en la Figura 1. A partir de estos plásmidos se determinó, en ambas cadenas del ADN y por el método de Sanger (*dideoxi*), la secuencia nucleotídica de 3617 bp entre los sitios de Sac I y Sma I (Figura 1, parte inferior). Estos datos se presentan en la SEQ N° 1. Mediante comparaciones de secuencia confirmamos que parte de la secuencia genómica determinada se corresponde con la del ADNc *Ha ds10* [Almoguera y Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; número de acceso en GenBank X59699]. La secuencia de aminoácidos de la proteína codificada por el gen *Ha ds10 G1* se indica bajo las secuencias nucleotídicas correspondientes. En el ADN genómico, la zona codificante está interrumpida por un intrón anómalamente largo (de 1024 bp), aunque situado en una posición conservada en otros genes LEA de clase I [ver datos revisados por

Simpson GC, Leader DJ, Brown JWS y Franklin T, en *Characteristics of Plant pre-mRNA Introns and Transposable Elements*; *Plant Mol. Biol. LabFax*, pp. 183-252; Croy RRD Ed., Bios Scientific Publishers Ltd. 1993]. La única diferencia,

entre las secuencias genómicas que codifican el ARNm y las del ADNc, fue una

5 inversión de dos nucleótidos (GC en vez de CG) dentro del segundo exón (en

las posiciones +1176 y +1177 desde el codón de iniciación); lo que provoca un

cambio de un aminoácido (S en vez de T) en la secuencia de la proteína. La

diferencia se debe a un error (debido a una compresión) en la lectura inicial de

las reacciones de secuencia del ADNc. Las secuencias de *Ha ds10 G1* que

10 hemos determinado incluyen también 1576 bp, del promotor del gen y

secuencias 5'-flanqueantes; y 553 bp de secuencias genómicas 3'-flanqueantes,

no presentes en el ADNc original.

Mediante la técnica de extensión del cebador (*primer extension*), se  
determinaron tres posibles sitios de iniciación de la transcripción en el promotor:

15 de *Ha ds10 G1*. Dos de estos sitios han sido confirmados mediante otras

técnicas (sitios 1 y 2, indicados por flechas en la SEQ N° 1). Para ello se utilizó,

según el procedimiento descrito por Domon *et al.* [Domon C, Evrard JL, Pillay

DTN, y Steinmetz A. *Mol. Gen. Genet.* 229:238-244, 1991], ARN total de

embriones de girasol hibridado con el cebador sintético: 5'-

20 CTCCTGTTCCGGAATTTTGCGTGT-3'; cuya secuencia corresponde a la de la

cadena no codificante de *Ha ds10 G1*, entre las posiciones +25 y +48, desde el

codón de iniciación. Las hibridaciones con el cebador se hicieron a 62°C. Los

híbridos se extendieron con transcriptasa reversa de AMV, durante 90 min. a

42°C. Los productos de extensión se analizaron en geles de secuenciación

25 PAGE al 6%, junto con reacciones de secuencia producidas usando el mismo

cebador. Los sitios de iniciación 1 y 2 (en las posiciones -33 y -25; ver SEQ N

1) son funcionales, y se detectan de forma independiente usando la técnica de

protección frente a la ribonucleasa A (RNasa A, ver Figura 3A). Un tercer sitio

de iniciación (sitio 3, en la posición -119 de la SEQ N 1) no fue confirmado

30 claramente, mediante dicha técnica. Estos sitios de iniciación delimitan

funcionalmente el extremo 3' del promotor del gen *Ha ds10 G1*.

El análisis de las secuencias proximales del promotor del gen *Ha ds10G1* mostró que dos de los sitios de iniciación detectados (los sitios 1 y 2) se encuentran a una distancia apropiada de una posible secuencia TATA (en la posición -86). El posible sitio mas distal (sitio 3, -119) no tiene secuencias TATA  
 5 claras situadas en su proximidad. Además de estos elementos del promotor, se observaron dos posibles "cajas" RY (RY1 e RY2 en las posiciones -129 y -65 de la SEQ N° 1), como las que participan en regulación de la expresión en semillas de numerosos genes de plantas [Dickinson CD, Evans RP, y Nielsen RC, en *Nucleic Acids Research* 16: 371, 1988 ].

10 Hemos modificado la caja RY1 situada en -129; verificando, mediante experimentos de expresión transitoria en embriones de girasol, su requerimiento funcional para la trans-activación del promotor de *Ha ds10G1* por factores transcripcionales de tipo ABI3 [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. Para ello, preparamos  
 15 modificaciones de las fusiones ds10::GUS construídas para estudios en plantas transgénicas (ver el Ejemplo 6.3 y la Figura 5). Los genes quiméricos contenidos en dos de estas fusiones (ds10F1y ds10F2) se purificaron como fragmentos de ADN que se subclonaron por ligación en el vector pBluescript SK+ (Promega); cambiando así las secuencias del vector binario por otras de menor tamaño,  
 20 mas útiles para realizar experimentos de expresión transitoria. Así, usando el fragmento Sal I - Eco RI (con el gen quimérico obtenido a partir de ds10F1), obtuvimos el plásmido pSKds10F1. En el caso de ds10F2, el fragmento de Sph I - Eco RI (desde la posición -125 en *Ha ds10 G1*, hasta el extremo 3' de *nos*) se ligó al fragmento complementario (que contiene el promotor y secuencias  
 25 5'flanqueantes de *Ha ds10 G1*), purificado tras la digestión de pSKds10F1 con Sph I y Eco RI, resultando en el plásmido pSKds10F2. Finalmente a partir de los plásmidos pSKds10F1 y pSKds10F2 (mapas no mostrados) se obtuvieron versiones mutagenizadas de los mismos tras la digestión de su ADN con Sph I, haciendo romos los extremos resultantes mediante tratamientos con ADN  
 30 polimerasa de T4, seguidos de re-ligación del ADN. De esta forma obtuvimos los plásmidos pSKds10F1ÆERY y pSKds10F2ÆERY (mapas no mostrados). Estos plásmidos difieren únicamente en una deleción de 5 nucleótidos entre las

posiciones -126 y -122 del promotor de *Ha ds10 G1*. Estos cambios destruyeron la caja RY1 presente en los genes quiméricos ds10F1 y ds10F2 (ver Figuras 1, 2 y 5), lo que se verificó mediante reacciones de secuenciación por el método de Sanger (dideoxy), usando el cebador 5'CTCCTGTTCCGGAATTTTGC GTGT3' (cadena no codificante de *Ha ds10G1*, entre las posiciones +25 y +48 ).

Los experimentos de trans-activación en expresión transitoria se realizaron mediante el bombardeo de embriones de girasol con proyectiles cubiertos de mezclas de ADN de distintos plásmidos. Estas mezclas contienen un plásmido de referencia, pDO432 [Ow DW, Wood KV, deLuca M, de Wet JR, Helinski D y Howell SH. Science 234: 856-859, 1996], con el gen de la luciferasa (LUC) de luciérnaga (*Photinus pyralis*) bajo el control del promotor CaMV 35S; la fusión ds10::GUS ensayada en cada caso (con las secuencias RY1 intactas o modificadas), y un plásmido efector, pABI3, que expresa el factor ABI3 bajo el control del promotor CaMV 35S. pABI3 se obtuvo sustituyendo el ADNc de Pv: ALF en el plásmido pALF [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995], por el ADNc de ABI-3. El ADNc de ABI3 se clonó como un fragmento Xba I (hecho romo con *klenow*) - Eco RI (parcial), purificado a partir del plásmido pcabi3-4F [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. El plásmido pABI3 se añade a la mezcla, o se omite, para probar el efecto del factor ABI3 sobre la expresión GUS de la fusión ensayada. Los experimentos se realizaron esencialmente como se describe por Bobb *et al.*, [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995], con las siguientes modificaciones. Los embriones de girasol (17-20 dpa) se prepararon como sigue. Las semillas de girasol se esterilizan con lavados en etanol 70% durante 1 min, y en 2% de hipoclorito sódico con una gota de Tritón X-100 durante 40 min, finalizados con varios lavados con agua destilada; tras los que se pelan en condiciones estériles. Los embriones se cortan longitudinalmente (separando sus dos cotiledones) y se colocan con la superficie cortada, sobre placas con medio sólido MS, que contiene 2% sacarosa y 0.5 M sorbitol. A continuación se precultivan durante 2-4 h en oscuridad y temperatura ambiente (25°C). Todos los plásmidos fueron purificados usando el *Quantum midiprep kit* (Biorad).

- Normalmente se usaron para cada disparo: 0.2 µg del plásmido de referencia, 1 µg del plásmido ds10::GUS y 1 µg del plásmido efector (o la misma cantidad del plásmido pJIT82 en los controles negativos). Para la preparación de las partículas de oro, así como la precipitación del ADN sobre las mismas, se siguió el método descrito por Chern *et al.* [Chern MS, Bobb AJ y Bustos M. *The Plant Cell* 8: 305-321, 1996]. El bombardeo de partículas se llevó a cabo con el sistema *Biolistic PDS-1000 He* (Biorad). Las condiciones de bombardeo fueron: Membrana de ruptura de 1550 psi, partículas de oro de 1.6 µm de diámetro, distancia de la membrana de ruptura al *macrocarrier* de 8 mm, distancia del *macrocarrier* a la rejilla de 6 mm; y distancia al tejido a bombardear de 6 cm. Los cotiledones bombardeados se incubaron durante 24 h a 28 °C en la oscuridad; tras lo cual se ensayó la actividad GUS (referida a la actividad LUC), como se describe por Bobb *et al.* [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995].
- La adición del plásmido efector pABI3 tuvo un efecto claro sobre la expresión relativa de GUS/LUC en bombardeos con la fusión pSKds10F2 (incremento medio de actividad relativa Å46.2X). En cambio, si la transactivación se hace con el mismo plásmido mutado en la caja RY1 (pSKds10F2ÆRY1), se observó un descenso significativo del incremento medio de actividad relativa debido al efecto de ABI3 (Å26.3 X). Este resultado, mostrado en la figura 2, confirma el requerimiento funcional de la secuencia RY1 (posición -129 en la SEQ N° 1). Por lo tanto esta caja RY participa en la activación transcripcional en semillas del promotor *Ha ds10 G1*, por factores del tipo ABI3 [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. Otras secuencias del promotor (por ej. RY2 en -65) también pudieran contribuir al efecto de transactivación observado, ya que la mutación ensayada no destruye completamente el efecto activador de ABI3.
- EJEMPLO 2: Acumulación y localización específica del mRNA *Ha ds10* en embriones de girasol:

Los patrones de acumulación de los ARN mensajeros del gen *Ha ds10G1* se determinaron mediante la técnica de la protección frente a la Ribonucleasa A (RNasa A), descrita con detalle por Almoguera *et al.* [Almoguera C, Coca MA, Jordano J. *Plant Physiol.* 107: 765-773, 1995]. Para ello, se utilizaron muestras

5 de ARN total preparadas a partir de embriones de semillas en distintos estados de desarrollo en condiciones normales de crecimiento [Almoguera y Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; Coca *et al.*, *Plant Mol. Biol.* 25: 479-492, 1994]; de germínulas de 3 días tras la imbibición (dpi); y de distintos órganos de plantas adultas antes de la floración. Los ARN de germínulas y plantas se

10 prepararon a partir de material vegetal obtenido tanto en condiciones de crecimiento controlado [Almoguera y Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, y Jordano J. *Plant Mol. Biol.* 25: 479-492, 1994; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996], como tras tratamientos de estrés: déficit de agua [Almoguera C, Coca

15 MA, y Jordano J. *Plant J.* 4: 947-958, 1993; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996]; o tras la adición de hormonas como el ácido abscísico [Almoguera C y Jordano J. *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996]. Las condiciones empleadas en cada tratamiento se describen con

20 detalle en las referencias citadas en cada caso. La ribosonda usada para detectar los ARNm de *Ha ds10 G1* tiene una longitud de 396 nucleótidos, de los cuáles 63 son secuencias del vector pBluescript SK+ y el resto la secuencia de la cadena no-codificante de *Ha ds10 G1* entre las posiciones +212 y -121 (Sph I). Esta sonda hibrida con el extremo 5' de los ARN mensajeros de *Ha ds10*

25 *G1*, sobrepasando el sitio mas distal de iniciación de la transcripción (sitio 3, SEQ N° 1), lo que permite detectar ARN mensajeros (ARNm) producidos a partir de los tres sitios de iniciación y la verificación experimental de las posiciones de iniciación. Esta ribosonda se preparó por transcripción *in vitro*, usando la ARN polimerasa T3, y como molde ADN del plásmido ds10G1S3Æ4.4. (Figura 1) que

30 contiene las secuencias de *Ha ds10G1* entre -1576 (Sal I) y +212, clonadas en el vector pBluescript SK+.

Los resultados en la Figura 3 muestran que los ARN mensajeros de *Ha ds10 G1* se detectan únicamente en semillas. Los niveles mayores de acumulación se observan en torno a 18-20 dpa, detectándose la expresión del gen a partir de los 10 dpa y desapareciendo tras la germinación (Figura 3). Los

5 tratamientos con ABA, o déficit de agua no indujeron la acumulación de los ARN mensajeros de *Ha ds10 G1* (datos mostrados para ABA en germinulas; Figura 3). Como control positivo en las muestras de ARN analizadas para los distintos tratamientos, realizamos hibridaciones (datos no mostrados) con otra ribosonda

10 de 651nucleótidos del gen *Ha hsp17.7 G4*, descrita anteriormente [Coca *et al.*, Plant Mol. Biol. 31: 863-876, 1996]; ya que dicho gen se expresa en respuesta a los distintos tratamientos ensayados. Estos análisis demostraron que los ARNm de *Ha ds10 G1* se acumulan exclusivamente en semillas, en condiciones normales del desarrollo y desde etapas tempranas de la maduración, confirmándose la iniciación a partir de al menos los sitios 1 y 2 (indicados en

15 SEQ N° 1). La banda indicada por el número 3 (Figura 3) no coincide bien con el tamaño esperado para el sitio de iniciación 3 (SEQ N°1). Esta banda pudiera deberse a la protección de secuencias de ARN mensajeros de un gen muy homólogo; o bien del mismo *Ha ds10 G1*, conteniendo secuencias del intrón (ARNm sin procesar).

20 La distribución de los ARNm de *Ha ds10 G1* en embriones de girasol, fue investigada mediante experimentos de localización por hibridación *in situ*. Para ello los embriones se incluyeron en parafina, fijaron, seccionaron, e hibridaron con sondas específicas; esencialmente como se describe por Molinier [en la

25 tesis: *Diplome D' Etudes Approfondies de Biologie Cellulaire et Moleculaire, Université Louis pasteur, Strasbourg, 1995*]. El tiempo de fijación se incrementó, desde 16 h a 4°C hasta 5 días, aumentando según la edad de los embriones. La deshidratación de los embriones fijados se hizo por incubaciones sucesivas (2 veces cada una durante 30-90 min.) en etanol al

30 10%,20%,30%,40%,50%,60%,70%,95%, y 100%; seguidas de inmersión en tolueno al 100% (1-3h, 2 veces). Los embriones fijados se incluyeron primero en tolueno:parafina (1:1), a 65°C durante 6-15 h, seguido de 5 inclusiones consecutivas en parafina, a 60°C durante 5-15 h. Las prehibridaciones e

hibridaciones con las sondas se hicieron a 45°C. La ribosonda específica de *Ha ds10 G1*, correspondiente al extremo 3'- del ARNm, se preparó como sigue. El plásmido ds10G1S1 (Figura 1) se usó como molde para preparar dos sondas por transcripción *in vitro* [Almoguera C, Coca MA y Jordano J. *Plant Physiol.* 107: 765-773, 1995] marcando con DIG-UTP. La sonda ds10-3'(-) se obtiene digiriendo el ADN del plásmido con Pvu II y efectuando la transcripción con ARN polimerasa T3. Esta sonda corresponde a la cadena no-codificante de *Ha ds10 G1* entre las posiciones +1202 (Pvu II en el segundo exón) y +1592 (extremo 3'). La segunda sonda [ds10-3' (+), usada como control], se preparó digiriendo el ADN de *Ha ds10 G1S1* con Bam HI (en el *polylinker*); y efectuando la transcripción con ARN polimerasa T7. La sonda ds10-3'(+) contiene la cadena codificante de *Ha ds10 G1*, entre las posiciones +870 y +1592. La especificidad de hibridación se determinó mediante experimentos de *Southern* similares a los descritos por Almoguera y Jordano [*Plant Mol. Biol.* 19: 781-792, 1992]. Mientras la hibridación con una sonda del ADNc completo detecta bandas correspondientes a unos 4-5 genes distintos en el genomio de girasol [Almoguera C, y Jordano J. *Plant Mol. Biol.* 19: 781-792, 1992]; usando la sonda ds10-3'(-) podemos detectar un único gen (con una ligera hibridación cruzada con otro; datos no mostrados).

Los resultados obtenidos en los experimentos de localización de ARN se muestran en la Figura 4. La sonda ds10-3'(-) es complementaria y de polaridad opuesta a los ARNm de *Ha ds10 G1*, lo que permite su detección. Los resultados obtenidos concuerdan con los datos de protección mostrados en la Figura 3, y muestran su acumulación en embriones desde los 12-15 dpa (Figura 4A) hasta los 21-28 dpa (Figuras 4C, F y H). Esta acumulación ocurre a niveles altos, lo que se deduce del corto tiempo preciso para su detección histoquímica (2-4 horas). En embriones inmaduros (Figura 4A) la distribución de los ARNm de *Ha ds10 G1* es homogénea y comparable (Figura 4B) a la del ARNr 18S, que se detecta usando otra ribosonda correspondiente al fragmento G (Eco RI) del gen 18S de rábano [descrito por Delcasso-Tremousaygue D, Grellet F, Panabieres F, Ananiev E D, y Delseny, M. En *Eur. J. Biochem.* 172: 767-776, 1988]. En embriones mas maduros (21 dpa, Figura 4C) los ARNm de *Ha ds10 G1* se

localizan también bastante homogéneamente, comenzando a detectarse una acumulación mas intensa en los haces vasculares (*procambium*), algo que no se observa con la sonda del ARNr 18S ni en éste ni en otros estadíos del desarrollo (Figuras 4D, B y G). Finalmente a los 28 dpa, los ARNm de *Ha ds10 G1* se localizan preferente en el parénquima en empalizada, un tejido especializado en la deposición de sustancias de reserva, situado en la cara interna de los cotiledones (Figuras 4F y H). Las localizaciones con la sonda ds10-3' (+), de la misma polaridad que los ARNm de *Ha ds10 G1*, no dieron señales de hibridación; lo que controló los experimentos descritos anteriormente (comparar las Figuras 4C y E). Estos experimentos demostraron que los patrones de expresión de los ARNm de *Ha ds10 G1* en girasol son muy especiales. La expresión observada en semillas, con altos niveles de acumulación desde etapas tempranas de la maduración embrionaria (10-12dpa), se combina con distribuciones espaciales que cambian desde la homogeneidad hasta la mayor abundancia en tejidos de deposición de sustancias de reserva (parénquima en empalizada). La distribución y patrones de acumulación de los ARNm de *Ha ds10 G1* es distinta a la que presentan otros genes vegetales pertenecientes a la misma familia [Wurtele ES, Wang HQ, Durgerian S, Nikolau BJ y Ulrich TH. *Plant Physiol.* 102: 303-312, 1993; Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., y Delseny, M., *Mol. Gen. Genet.*, 238: 409-418, 1993]. Estos resultados indican la posible utilidad, para la modificación de semillas por Ingeniería genética, de genes quiméricos que incorporen las secuencias reguladoras de *Ha ds10 G1*.

### 25 EJEMPLO 3: Construcción de genes quiméricos ds10G1::GUS y su análisis en plantas transgénicas de tabaco:

Como ejemplo para los posibles usos del promotor y las secuencias reguladoras del gen *Ha ds10 G1*, en la construcción de genes quiméricos con expresión específica en semillas de plantas transgénicas, describimos a continuación la construcción y el análisis en plantas transgénicas de tabaco de 4 fusiones traduccionales ds10G1::GUS (Figura 5). Dichas fusiones contienen, para su análisis funcional, el promotor y distintas combinaciones de secuencias

flanqueantes e intragénicas del gen *Ha ds10 G1*. Estas 4 fusiones proporcionan elevados niveles de expresión del gen indicador (GUS) en semillas desde etapas tempranas de la maduración (Figura 6), confirmando nuestras observaciones en el sistema homólogo (Ejemplo 2, Figuras 1-4).

- 5           La primera de estas construcciones, ds10F1 (Figura 5) se obtuvo a partir del plásmido ds10G1S3 (Figura 1), que contiene las secuencias genómicas de *Ha ds10 G1* entre Sal I (-1576) y Eco RI (+1086), subclonadas en los sitios de restricción correspondientes del vector pBluescript SK+ (Promega). Mediante tratamiento con Exonucleasa III del ADN de ds10G1S3 (previamente digerido con Hind III y Pst I), se delecionaron las secuencias de *Ha ds10 G1* entre Eco RI (+1086) y la posición +98 (en el primer exón), dando lugar al plásmido ds10G1S3Æ10.5 (Figura 1). Dicho plásmido se digirió con Bam HI (diana de restricción del *polylinker* situada inmediatamente adjacente a la posición +98 de *Ha ds10 G1*), rellenándose a continuación los extremos del ADN digerido usando el fragmento de Klenow de la ADN polimerasa I. A continuación el ADN se digirió con Sal I, purificándose el fragmento de 1679 p.b. que contiene las secuencias de *Ha ds10 G1* entre Sal I (-1576) y el extremo relleno de Bam HI. Este fragmento se clonó entre los sitios de Sal I y Sma I del vector binario pBI 101.2, resultando en ds10F1, una fusión traduccional que contiene 1576 nucleótidos de secuencias 5'-flanqueantes de *Ha ds10 G1* (desde el ATG) y los primeros 98 nucleótidos de la zona codificante, en fase con el gen GUS (Figura 5). La fusión ds10F2 se derivó a partir de ds10F1 mediante la inserción de un fragmento de ADN genómico de *Ha ds10G1* comprendido entre las posiciones (Figura 1) de +1205 (Pvu II), y Eco RI (+4670). Dicho fragmento contiene parte del segundo exón y 3370 nucleótidos de secuencias 3'-flanqueantes (a partir de codón de terminación en la posición +1301); y reemplaza a las secuencias nos-3' en la fusión ds10F1. El inserto Pvu II- Eco RI se purificó a partir de ADN del plásmido ds10G1S2 (Figura 1). Para la inserción de dicho fragmento, el ADN de ds10F1 se digirió con Sac I y los extremos del ADN se hicieron romos mediante tratamiento con la ADN polimerasa I de T4. A continuación, el ADN así tratado se digirió con Eco RI, purificándose el fragmento con las secuencias de *Ha ds10G1*. Este fragmento se ligó al inserto Pvu II- Eco RI anteriormente

descrito (con las secuencias 3'-flanqueantes de *Ha ds10 G1*), resultando en la fusión ds10F2 (Figura 4). La fusión ds10F2Æ (Figura 4) se obtuvo a partir de ds10F2, mediante la delección de las secuencias 3'-flanqueantes de *Ha ds10G1* entre Xba I (Å+2830) y Eco RI (Å+4670). Para ello, el ADN de ds10F2 se digirió con ambos enzimas; religándose, tras hacer romos los extremos de ADN resultantes con el fragmento de Klenow de la ADN polimerasa I. Finalmente, la cuarta fusión (ds10F3, Figura 5) se obtuvo a partir de un fragmento de ADN genómico de *Ha ds10 G1* entre Sal I (-1576) y Pvu II (+1204), purificado a partir del plásmido ds10G1S6 (Figura 1) tras la digestión con ambas enzimas de restricción. Este fragmento se ligó con ADN del vector pBI101.3, digerido previamente con Sal I y Sma I. La fusión ds10F3 contiene de esta forma el promotor y las mismas secuencias 5'-flanqueantes de *Ha ds10 G1* presentes en la fusión ds10F1, así como el primer exón (de +1 a +145), el intrón completo (de +146 a +1169) y parte del segundo exón de *Ha ds10 G1* (de +1170 a +1204), fusionado en fase con el gen GUS de pBI 101.3. En todos los casos la secuencia de nucleótidos correspondiente a la zona de fusión, entre las secuencias GUS y las de *Ha ds10 G1*, se comprobó mediante reacciones de secuenciación con el método de Sanger (dideoxi), usando como cebador las secuencias de GUS: 5'-ACGCGCTTTCCCACCAACGCTG-3'.

El ADN-T en las fusiones ds10F1, ds10F2, ds10F2Æ, y ds10F3 (Figura 5) fue movilizado desde *A. Tumefaciens* (LBA 4404), obteniéndose distintas plantas transgénicas de tabaco con integraciones independientes de cada gen quimérico. Estas plantas fueron obtenidas y caracterizadas mediante procedimientos estándar que se describen con detalle por Coca MA, Almoguera C, Thomas TL y Jordano J, [en *Plant Molecular Biology*, 31: 863-876, 1996]. En dichas plantas, la expresión del gen *GUS* se analizó tanto en semillas en desarrollo en condiciones normales de crecimiento (sin estrés exógeno); como en tejidos de germínulas, investigándose en este último caso los cambios de expresión inducidos por tratamientos con ABA y de desecación. Los análisis de semillas se realizaron con las plantas transgénicas originales (T0); mientras que para los de germínulas se utilizaron descendientes de estas plantas (T1), segregantes para los genes quiméricos. Se hicieron tanto estudios cuantitativos,

mediante el análisis fluorimétrico de los niveles de expresión *GUS* y de sus patrones temporales, como estudios cualitativos analizando histoquímicamente los patrones espaciales de expresión (especificidad de tejido). Estos estudios se hicieron como se describe con detalle por Coca MA, Almoguera C, Thomas TL y Jordano J, [en *Plant Molecular Biology*, 31: 863-876, 1996]. En total se obtuvieron y analizaron los siguientes números (entre paréntesis) de plantas transgénicas de tabaco, T0 "funcionales", con los genes quiméricos ds10F1 (14), ds10F2 (7), ds10F2Æ (8) y F3 (23). Estas plantas mostraron elevados niveles expresión del gen *GUS* en semilla (como consecuencia de la actividad del promotor y secuencias reguladoras del gen *Ha ds10 G1*), según se ilustra en la Figura 6 (paneles A-C). La integración de los distintos genes quiméricos en el ADN de las plantas transgénicas fue caracterizada mediante *Southerns* genómicos usando sondas de la región codificante de gen *GUS*; amplificaciones PCR de las secuencias próximas al empalme ds10::GUS, usando los cebadores 5'-ACGCGCTTTCCCACCAACGCTG-3' (GUS) y 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI); o mediante test de segregación de la resistencia a la Kanamicina (conferida por el gen *nptII*), relizados según se describe en [Jordano J, Almoguera C, y Thomas TL, *The Plant Cell* 1: 855-866, 1989]. Estos análisis determinaron que las plantas T0 seleccionadas para los estudios de expresión en semillas contenían de 1 a 5 integraciones independientes del gen quimérico correspondiente. La Figura 6 (adjuntada con esta solicitud) ilustra los resultados mas relevantes, obtenidos en el estudio de la expresión en plantas transgénicas de los genes quiméricos analizados. Estos resultados se describen con detalle a continuación.

La expresión *GUS* durante la maduración de las semillas en condiciones de crecimiento controladas (sin estrés exógeno), se analizó mediante ensayos fluorimétricos (Figura 6A) e histoquímicos (resumen en Figuras 6B-E). Los ensayos fluorimétricos se realizaron con semillas en estadíos definidos de maduración, a los 12, 16, 20, 24 y 28 días post-anthesis (dpa). Por cada planta T0 y estadio de maduración se preparon extractos de dos cápsulas florales distintas, y se ensayó la actividad *GUS* con Methilumbeliferilglucoronido (MUG) por duplicado (en total cuatro determinaciones de actividad por estadio de

desarrollo y planta transgénica individual). La significación estadística de las diferencias observadas con las distintas fusiones GUS se determinó, tras la normalización logarítmica de los datos obtenidos, mediante análisis de la varianza [ANOVA, ver: Nap JP, Keizer P, y Jansen R, en *Plant Molecular Biology Reporter* 11: 156-164, 1993]. Los ensayos histoquímicos se hicieron con material diseccionado a partir de semillas, en estadios de desarrollo definidos, procedentes de los siguiente números de plantas transgénicas: d10F1, 5; ds10F2, 6; ds10F2Æ, 6; y dsF3, 19. El endospermo y los embriones diseccionados a partir de semillas individuales se tiñeron con X-gluc, durante 150 min a 25°C, analizándose de esta forma aproximadamente 150 semillas de cada planta transgénica.

Todos los genes quiméricos produjeron niveles elevados de expresión GUS en semillas, alcanzándose valores máximos medios de  $1.65 \times 10^6$  pmol MU/ mg x min (Figura 6A: a los 24 dpa). Los ensayos histoquímicos confirmaron estos altos valores de actividad, ya que tanto los embriones (Figuras 6B y C) como el endospermo (Figura 6C) se tiñeron fuertemente a partir de los 12 dpa (Figura 6B), y con sólo 150 min de reacción. En ambos casos se observaron distribuciones espaciales de la actividad GUS bastante homogéneos (Figura 6B-C). Además, estos patrones de expresión no difirieron cualitativamente entre las plantas transgénicas de los distintos genes quiméricos (datos no mostrados).

Los ensayos fluorimétricos revelaron interesantes diferencias cuantitativas entre las distintas fusiones ds10::GUS. Estas diferencias dependen de las secuencias de *Ha ds10 G1* presentes en las fusiones. En algunos casos se ha podido mostrar la significación estadística de estas diferencias (con un nivel de confianza del 95%), lo que demuestra experimentalmente la contribución de las distintas secuencias ensayadas (promotor y secuencias 5'-flanqueantes, secuencias codificantes, 3'-flanqueantes, y del intrón) a los patrones de expresión embrionaria observados. La presencia en las fusiones de secuencias 3'-flanqueantes de *Ha ds10 G1* incrementa los niveles de expresión GUS en semillas entre 20 y 28 dpa (comparar las fusiones ds10F2 y ds10F2Æ, con ds10F1 en las Figuras 5 y 6A). Esta diferencia es estadísticamente significativa (por ejemplo a 28 dpa:  $F= 5.397$ ,  $P=0.0213$ ), y está causada por las

secuencias de *Ha ds10 G1* presentes en la fusión ds10F2Æ (ver Figura 5); ya que no se encontraron diferencias significativas entre la actividad GUS de ds10F2 y ds10F2Æ (por ejemplo, también a los 28 dpa,  $F=0.274$ ,  $P=0.6015$ ; ver Figura 6A). En el caso de ds10F2Æ, el efecto estimulador de las secuencias 3'-

5    flanqueantes también se produce, y es altamente significativo, en etapas mas tempranas de la maduración embrionaria (Figura 6A, 16 dpa;  $F=16.607$ ,  $P=0.001$ ). En cambio, en estas etapas (entre 12 y 16 dpa) las actividades GUS de ds10F1 y ds10F2 no difieren significativamente entre sí (por ejemplo, a 16 dpa:  $F=2.762$ ,  $P=0.0983$ ; ver Figura 6A). En conjunto estos resultados muestran

10    que ds10F2Æ es la fusión construida y ensayada que funciona mejor en semillas de tabaco desde los 16dpa; y que esto se debe al efecto de las secuencias 3'-flanqueantes de *Ha ds10 G1* incluidas en ella. Desconocemos si este efecto se produce por mecanismos de activación transcripcional, estabilización de ARNm, o por combinación de ambos tipos de mecanismos. En

15    cualquier caso el efecto es claro, y de posible utilidad para diseñar nuevos genes quiméricos de expresión mas eficiente en semillas, desde etapas relativamente tempranas de la maduración embrionaria (veáse también el apartado de "Otros Ejemplos").

Por otra parte, la comparación entre las actividades GUS de las plantas

20    con las fusiones ds10F1 y ds10F3 nos permitió investigar los posibles efectos de la presencia del intrón (y/o de las secuencias codificantes de *Ha ds10 G1* en las que difieren estas fusiones, Figura 5) sobre la expresión de ambas. En semillas de tabaco transgénico estas comparaciones demuestran que la presencia del intrón (mas el primer exón completo y parte del segundo exón) no

25    tiene efectos positivos sobre la expresión GUS, que por lo tanto debe de estar básicamente conferidas por el promotor y secuencias de *Ha ds10 G1* presentes en ds10F1 (Figura 6A). Así por ejemplo, las actividades de ds10F1 y ds10F3 no difieren estadísticamente entre 12 y 28 dpa, salvo a los 20 dpa ( $F=4.73$ ,  $P=0.031$ ), y entonces la presencia de las secuencias adicionales en ds10F3

30    redujo significativamente la actividad GUS observada. Por lo tanto, aunque es altamente probable que el intrón se procese correctamente en semillas de sistemas heterólogos como el tabaco (carecemos de una prueba formal de ello),

su posible papel regulador en el desarrollo embrionario no está claro. Sin embargo otras observaciones no excluyen que las secuencias adicionales de *Ha ds10 G1* en *ds10F3* (incluyendo el intrón) puedan tener papeles reguladores en otros tejidos (ver, mas adelante, el efecto de éstas secuencias sobre la expresión residual de las fusiones *ds10::GUS* en el polen y en germinulas).

La especificidad embrionaria (en semillas) de la expresión GUS conferida por las secuencias *Ha ds10 G1* en plantas transgénicas de tabaco se investigó verificándola en otros tejidos; tanto en ausencia de estrés, como tras tratamientos de desecación o con ABA. En el caso de las plantas T0, el único tejido en el que, tanto mediante ensayos fluorimétricos como histoquímicos, se detectó actividad GUS fue en el polen maduro. En otros tejidos las actividades detectadas apenas superaron las del fondo (plantas de tabaco no transformadas). Por ejemplo, en hojas de plantas T0 de unos dos meses de edad: 0-50 pmol MU/ mg x min. Las actividades detectadas en polen son marginales (casi tres órdenes de magnitud inferiores) comparadas con las de semillas de las mismas plantas transgénicas. Además dicha expresión pudiera ser artefactual y depender del uso, como indicador, del gen GUS en las fusiones [según Uknes S, Dincher S, Friedrich L, Negrotto D, Williams S, Thompson-Taylor H, Potter S, Ward E, y Ryals J, en *the Plant Cell* 5: 159-169, 1993]. Sin embargo, de forma sorprendente, observamos que la actividad medida en el polen de 9 plantas *ds10F3* fue ( $136 \pm 64$  pmol MU/ mg x min) significativamente inferior a la de 5 plantas *ds10F1* ( $6427 \pm 1294$  pmol MU/ mg x min;  $F = 72.573$ ,  $P = 0.0001$ ). Esto último pudiera indicar que, a diferencia de lo que ocurre en semillas durante la mayor parte de la maduración del embrión (Figura 6A), la presencia de las secuencias adicionales de *Ha ds10 G1* en *ds10F3* (incluyendo el intrón) pudiera reducir la expresión, de genes quiméricos que las contengan, en otros tejidos o momentos del desarrollo.

Adicionalmente, se verificó si la expresión de las fusiones *ds10::GUS* puede inducirse por hormonas (ABA) o tratamientos de estrés (déficit de agua) en plantas transgénicas (T1) de tabaco en distintos momentos de su ciclo vegetativo. Para ello seleccionamos, tras germinación en medio MS con 300 µg/ml de kanamicina, descendientes de 8 plantas originales distintas

conteniendo las fusiones ds10F1, ds10F2Æ y ds10F3; y otras 6 con ds10F2. Las germinulas resistentes se transplantaron a medio MS. Se realizaron distintos experimentos con germinulas, tanto a 8 como a 15 días tras la imbibición. Para los tratamientos con ABA, las germinulas se transplantaron a

5 placas de MS suplementadas con 100 µM ABA y se cultivaron en dicho medio durante 4 días a 25 °C y con iluminación. Las germinulas control también se transplantaron a medio MS sin ABA. El estrés hídrico se provocó colocando a las germinulas durante unas 5-6 horas dentro de una cabina de flujo entre dos papeles de filtro. Tras los distintos tratamientos, las germinulas se procesaron

10 bien individualmente (para los ensayos histoquímicos con X-gluc, mediante incubaciones de 14 h a 25 °C); o conjuntamente (*pool analysis*), para los ensayos fluorimétricos de la actividad GUS, realizados como se ha descrito anteriormente. Los tratamientos de plantas transgénicas adultas, se hicieron usando plantas individuales propagadas como clones vegetativos obtenidos de

15 cada planta original. Para ello, las germinulas seleccionadas de cada planta transgénica se transplantaron a vermiculita embebida con medio Hoagland 0.5X. De cada germinula se obtuvieron tres explantes completos, que tras recuperarse se pusieron en cultivo hidropónico en medio Hoagland líquido (0.5X). Los experimentos se realizaron cuando las plantas se habían recuperado por

20 completo del proceso de propagación, y tenían raíz, tallo y unas 10-12 hojas. Por lo tanto, para los distintos tratamientos se usaron plantas idénticas genéticamente procedentes de cada germinula transgénica seleccionada. Los tratamientos con ABA se hicieron añadiendo la hormona al medio (100 µM), analizándose la actividad GUS en las plantas a las 24h. El estrés hídrico se

25 indujo retirando la raíz del contenedor con el medio, analizándose igualmente las plantas a las 24h tras iniciar el tratamiento. El efecto de los distintos tratamientos se analizó en tres experimentos independientes realizados con los siguientes números de plantas T1 para cada fusión (entre paréntesis el número de plantas T0 de las que proceden en cada caso): ds10F1, 11 (6); ds10F2, 10

30 (5); ds10F2Æ, 5 (3); y ds10F3, 10 (5).

Los experimentos realizados tanto con germinulas como con plantas adultas confirmaron la especificidad embrionaria de la expresión conferida por

las secuencias de *Ha ds10 G1* a las distintas fusiones; aportando además indicios adicionales sobre el posible papel regulador de las secuencias de *Ha ds10 G1* presentes en ds10F3 (incluyendo el intrón) mencionadas anteriormente. Así, tanto en plantas adultas control, como tratadas, se detectaron actividades GUS mínimas (de 3 a 300 pmol MU/ mg x min) en todos los tejidos analizados (raíz, tallo, hojas y meristemo apical). Estos niveles de actividad están ligeramente por encima de los valores de fondo y pueden detectarse sólo fluorimetricamente (datos no mostrados).

En germinulas de 8 dpi la expresión de todas las fusiones es unos dos órdenes de magnitud inferior a los valores máximos alcanzados en semillas. Esta expresión decrece rápidamente entre los 8 y 15 dpi (por ej. ds10F1 pasa de  $2864 \pm 182$  a  $813 \pm 104$  pmol MU/ mg x min); y se restringe exclusivamente en los tejidos embrionarios (cotiledones), sin detectarse en otros tejidos vegetativos (radícula, hipocótilo, hojas) diferenciados tras la germinación (Figuras 6D y E, y datos no mostrados para las otras fusiones). Estos resultados confirman, en plantas transgénicas de tabaco, la especificidad embrionaria de la regulación por secuencias de *Ha ds10 G1*. Además de la reducción general de los valores de actividad GUS mencionada anteriormente, se observaron diferencias entre los valores de las distintas fusiones, algunas de ellas estadísticamente significativas. Estas diferencias son similares cualitativamente a las observadas en semilla (Figura 6A). Entre ellas, y por su posible interés aplicado, ilustramos la reducción de la expresión tras la germinación, mediada por las secuencias de *Ha ds10 G1* presentes en ds10F3 (incluyendo el intrón). Este efecto se observa, como una reducción significativa de actividad GUS al comparar los patrones de expresión de plantas ds10F1 y ds10F3 (Figuras 6D y E). El análisis estadístico de los datos cuantitativos de ds10F1 y ds10 F3 confirmó la significancia de esta diferencia, tanto a los 8 dpi ( $F = 4.36$ ,  $P = 0.04$ ) como a los 15 dpi ( $F = 4.39$ ,  $P = 0.039$ ). Adicionalmente, con las germinulas de ds10F1 se observó a los 8dpi una moderada inducción de GUS por los tratamientos con ABA que es estadísticamente significativa (de  $2864 \pm 182$  a  $5790 \pm 733$  pmol MU/ mg x min;  $F = 5.413$ ,  $P = 0.023$ ). En el caso de ds10F3 no hubo inducción significativa por el mismo tratamiento (de  $1502 \pm 195$  a  $2338 \pm 211$  pmol MU/ mg x min;  $F = 2.58$ ,  $P =$

0.11). Los distintos tratamientos no afectaron substancialmente la especificidad de tejidos, o el orden de magnitud de la expresión observada para las distintas fusiones *ds10::GUS* (datos no mostrados).

## 5 OTROS EJEMPLOS:

Igualmente pueden obtenerse, de forma análoga a la descrita con detalle en el ejemplo anterior, otros genes quiméricos que contengan secuencias 5'-flanqueantes, y(o) 3'-flanqueantes (terminadores), y(o) codificantes, procedentes *del Ha ds10 G1*, combinadas con secuencias procedentes de otros genes. Estos ejemplos no suponen complicaciones técnicas adicionales a los descritos con mas detalle en los apartados anteriores, por lo que son fácilmente realizables por personas con conocimientos suficientes en el sector de la técnica de la invención. Así por ejemplo, en las fusiones *ds10::GUS*, las secuencias *Ha ds10 G1* pudieran haber incluido otras secuencias 5'-flanqueantes (Figura 1) mas largas del mismo gen para aumentar su nivel de expresión en semillas, como describimos por ejemplo en [Coca MA, Almoguera C, Thomas TL, y Jordano J, en *Plant Molecular Biology* , 31: 863-876, 1996]. Igualmente, las secuencias GUS podrían ser substituidas por otras que codifiquen distintas proteínas o péptidos (naturales o artificiales), cuya producción regulada en semillas de plantas pudiera ser de interés industrial. Ejemplos de estas últimas posibilidades, dados de forma no exclusiva, serían la fusión a secuencias de *Ha ds10 G1* de secuencias codificantes de genes implicados en la biosíntesis de ácidos grasos en semillas [Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE y Davies HM, en *Science*, 257:72-74, 1992], de proteínas de reserva con composiciones ricas en determinados aminoácidos [Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, y Muntz K, en *Molecular and General Genetics* 242: 226-236, 1994], o de péptidos con actividades antigénicas o farmacológicas [Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, De Block M, Vandewiele M, De Clercq, Leemans J Van Montagu, M y Krebbers E, en *BioTechnology* 7: 929-932, 1989]. Estas fusiones se realizarían y utilizarían de forma análoga a como se describe en las publicaciones citadas a título de ejemplo (datos de forma no excluyente)

en cada caso. Para facilitar estas posibilidades, hemos construido un plásmido (ds10EC1) que contiene una cassette de expresión que incluye el promotor y las secuencias 5'- y 3'- flanqueantes de *Ha ds10 G1* presentes en ds10F2Æ (ver Figura 5). Entre ambas secuencias y mediante mutagénesis dirigida [Chen E y Przybila AE, en *BioTechniques* 17: 657-659, 1994] hemos añadido un sitio de restricción de Eco RI, que permite la inserción de secuencias de genes, o correspondientes a péptidos, como los mencionados anteriormente (disponibles en otros laboratorios, o que pudieran diseñarse o sintetizarse). El plásmido ds10EC1 se construyó a partir de ds10G1S3Æ10.5 (Figura 1). A partir de dicho plásmido, amplificamos por PCR las secuencias de *Ha ds10 G1* entre las posiciones -1574 (Sal I) y +98; usando ADN polimerasa Pfu y los cebadores 5'-ATTAACCCTCACTAAAG-3' (T3) y 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI). En este último los tres cambios de secuencia (señalados en minúscula) introducen el nuevo sitio de Eco RI en la posición del codón de iniciación. Tras la PCR se purifica un fragmento de ADN de 199 pb (*megaprimer*), que junto con el cebador 5'-AATACGACTCACTATAG-3' (T7) se usa para una segunda amplificación por PCR de ds10G1S3Æ10.5. El ADN amplificado (795 pb) se digirió con Eco RI y Sph I. El fragmento de ADN resultante (125 pb), con las secuencias de *Ha ds10 G1* entre Sph I (-126) y el nuevo sitio de Eco RI, se purificó y ligó; reemplazando en ds10G1S3 las secuencias de *Ha ds10 G1* (Figura 1) entre las posiciones -126 (Sph I) y 1086 (Eco RI). Tras este paso, la secuencia amplificada por PCR se verificó mediante secuenciación (método de Sanger) usando el cebador T3. Finalmente se insertó en el plásmido obtenido en el paso anterior un fragmento del ADN genómico de *Ha ds10 G1* (Figura 1), con secuencias entre +1086 (Eco RI) y +3000 (Xba I), obteniéndose la cassette ds10EC1 (Figura 4), clonada en el plásmido pBluescript SK+. El extremo 3' del ADN de ds10EC1 difiere del de ds10F2Æ únicamente en 119 nucleótidos adicionales, correspondientes a secuencias del intrón y del segundo exón de *Ha ds10 G1*. Además, las secuencias de *Ha ds10 G1* en ds10EC1 difieren de las correspondientes en ds10F2Æ en la ausencia de los nucleótidos 1-98 del primer exón (Figura 5).

Dado que la presencia de secuencias adicionales de *Ha ds10 G1* en ds10F3 (incluyendo el intrón, el primer exón y parte del segundo exón) redujo la expresión de este gen quimérico específicamente en tejidos no embrionarios (Ejemplo 3, Figuras 6D-E), es concebible que dichas secuencias pudieran utilizarse para conferir especificidad de semillas a otros genes quiméricos con distintos promotores. El diseño de dichos genes quiméricos no ofrece dificultades técnicas adicionales a las descritas en apartados anteriores: ver por ejemplo los procedimientos detallados para el uso de intrones de plantas con el fin de impedir la expresión de genes quiméricos en *Agrobacterium* [Mankin SL, Allen GC y Thompson WF. *Plant Molecular Biology Reporter* 15: 186-196, 1997]

Los genes quiméricos que contengan secuencias reguladoras de *Ha ds10G1* podrían ser transformados a otras plantas distintas de tabaco (el sistema modelo usado en el ejemplo 3). Entre las mismas hay plantas de gran importancia económica como por ejemplo: el girasol, la soja, la colza, la "canola", el maíz, el trigo, la cebada, el arroz, la "casava", la judía, el cacahuete, etc; cuya transformación genética es posible y está documentada suficientemente en la literatura científica: veáse por ejemplo Lindsey K, Ed. (1993). [*Plant Tissue Culture Manual*. Kluwer Academic Publishers]; y la revisión por Christou [*Trends in Plant Science*. 1: 423- 431, 1996]. Los resultados mostrados en el ejemplo 3 demuestran que, en tabaco, los genes contruidos con secuencias reguladoras de *Ha ds10 G1* tienen una elevada actividad desde etapas relativamente tempranas de la maduración embrionaria, manteniendo además la especificidad de semillas característica de la expresión de *Ha ds10 G1* en girasol. Estos resultados podrían obtenerse también con otras plantas, como las mencionadas anteriormente.

#### DESCRIPCIÓN DE LAS FIGURAS:

**Figura 1.** Parte superior: mapa de restricción de las secuencias genómicas de *Ha ds10 G1* que flanquean a su región codificante. Las líneas continuas sobre el mapa indican los distintos fragmentos de ADN genómico que han sido subclonados en el vector pBluescript SK+ (los nombres de los plásmidos respectivos se indican sobre cada fragmento). Los plásmidos preparados mediante delecciones con Exo III se indican sobre el plásmido de

partida (ds10G1S3ÆSacl), indicando en cada caso el extremo de la delección. En la parte inferior de la figura se incluye un mapa de restricción detallado de la región cuya secuencia nucleotídica ha sido determinada. La extensión de las distintas reacciones, usadas para ensamblar las secuencias de ambas cadenas de ADN, se indican mediante flechas horizontales (sobre el mapa para la cadena codificante, y bajo el mapa para la no-codificante). Los sitios de iniciación de la transcripción se indican por flechas. Se incluyen barras de escala para ambos mapas.

- Figura 2.** Implicación funcional de las secuencias RY1 (-129) en la trans-activación del promotor *Ha ds10 G1*. Experimentos de expresión transitoria realizados tras el bombardeo de embriones de girasol con micro-proyectiles cubiertos de ADN. Se representan los resultados de 5 experimentos independientes en los que las distintas mezclas de plásmidos (detalladas en el Ejemplo 1) se bombardearon por quintuplicado en cada experimento. Se representan las medias de las actividades  $\beta$ -glucoronidasa (GUS) normalizadas con la actividad luciferasa (LUC), así como los errores *standard* (indicados por barras). Clave: F2, pSKds10F2; F2ÆERY1, pSKds10F2ÆERY1; ABI3, muestras con el plásmido efector. Se aprecia una disminución significativa de la actividad relativa GUS/LUC, consecuencia de la mutación en la caja RY1. Las actividades basales para pSKds10F2 (sin incluir el plásmido efector) son del orden de  $46 \pm 8$ .

- Figura 3.** Patrones de acumulación de los ARNm del gen *Ha ds10 G1* en girasol. La autoradiografía mostrada corresponde a ensayos de protección frente a la RNasa A, tras hibridar una ribosonda del gen con distintas muestras de ARN total. Se observa la acumulación de mensajeros producidos a partir de los sitios de iniciación de la transcripción de *Ha ds10 G1* (como fragmentos protegidos indicados por las flechas numeradas). Estos fragmentos se detectan sólo en embriones (Emb) desde 10 a 20 dpa, y en semillas maduras (25 dpa); pero no en otras muestras analizadas, como germinulas (Germ) o germinulas tratadas con ABA (Germ + ABA). El carril tRNA corresponde a hibridaciones control con ARN t de levadura. Se indican con números y flechas las bandas

correspondientes a los mRNAs producidos a partir de los distintos sitios se iniciación. El sitio de iniciación número 3 (indicado entre paréntesis) no ha sido confirmado experimentalmente mediante *primer extension*. En el margen izquierdo se incluyen marcadores moleculares de tamaño (pBR322/Hpa III).

5

**Figura 4.** Localización de ARNm en secciones de embriones de girasol a los 12 (A y B), 21 (C-E), y 28 dpa (F-H). En cada caso se usaron las siguientes ribosondas : ds10 (-), A, C, F, H; ds10 (+), E, y 18S ARNr, B, D, G. Barras de escala = 500  $\mu$ m (Savo en F, 125  $\mu$ m). Parénquima en empalizada= pp. Las flechas señalan el *procambium*.

**Figura 5.** Mapas de restricción de las fusiones ds10::GUS y de la *cassette* de expresión optimizada ds10 EC1, construídas en los Ejemplos 3 y 4. Mediante distintos sombreados se indican las secuencias de *Ha ds10 G1* y de otros genes contenidas en cada caso. Los sitios de iniciación de la transcripción a partir del promotor de *Ha ds10 G1* están indicados por flechas.

**Figura 6.** Expresión de las fusiones ds10::GUS en semillas de plantas transgénicas de tabaco. Panel A: Compendio de todos los datos cuantitativos (determinaciones fluorimétricas). Se muestra el promedio de las actividades GUS observadas en semillas de las plantas transgénicas (T0) y su evolución en distintos momentos del desarrollo embrionario. Los datos correspondientes a cada fusión se indican mediante los símbolos en el inserto de la parte superior izquierda. Las barras indican los errores estándar. Paneles B-E: selección representativa con resultados de los experimentos de localización histoquímica de la actividad GUS: B.- embriones a los 12 dpa (plantas ds10F2Æ, T0). C.- embriones y endospermo a los 16 dpa (plantas ds10F2Æ, T0). D.- germínulas a los 15 dpi en condiciones control (plantas ds10F1, T1). E.- germínulas a los 15 dpi en condiciones control (plantas ds10F3, T1). En los paneles D y E, las flechas señalan los tejidos vegetativos sin actividad GUS (hojas e hipocótilo).

#### LISTA DE SECUENCIAS:

**SEQ N° 1:** Secuencia nucleotídica del gen *Ha ds10 G1*. Con flechas se indican los sitios de iniciación de la transcripción determinados experimentalmente (el sitio 3, que no ha sido confirmado mediante *primer extension* se señala entre paréntesis). La zona codificante se muestra mediante su traducción en aminoácidos, indicados mediante el código de una letra bajo la secuencia nucleotídica. El codón de terminación se indica por un asterisco. La secuencia está numerada (en el margen izquierdo) a partir del codón de iniciación. Las secuencias del intrón se indican mediante letras minúsculas. Las cajas TATA (en la posición -86) y RY(-129 y -65) mencionadas en el texto (Ejemplo 1) se señalan mediante subrayados.

⋮⋮⋮  
⋮⋮⋮

⋮⋮  
⋮⋮⋮

⋮⋮⋮  
⋮⋮⋮  
⋮⋮⋮  
⋮⋮⋮

⋮⋮⋮  
⋮⋮⋮  
⋮⋮⋮  
⋮⋮⋮

## LISTA DE SECUENCIAS

<110> Consejo Superior de Investigaciones Cientificas  
 <120> Promotor y secuencias reguladoras de Ha ds10G1: un gen LEA  
 expresado exclusivamente en semillas  
 <130> Ha ds10G1  
 <140>  
 <141>  
 <160> 1

<170> PatentIn Ver. 2.1

<210> 1  
 <211> 3617  
 <212> ADN  
 <213> Helianthus annuus

<220>  
 <221> exon  
 <222> (1577)..(1721)

<220>  
 <221> exon  
 <222> (2746)..(2879)

<220>  
 <221> intron  
 <222> (1722)..(2745)

<220>  
 <221> misc\_feature  
 <222> (1448)..(1455)  
 <223> RY box

<220>  
 <221> misc\_feature  
 <222> (1512)..(1519)  
 <223> RY box

<220>  
 <221> polyA\_signal  
 <222> (3139)..(3144)

<220>  
 <221> polyA\_signal  
 <222> (3220)..(3225)

<400> 1  
 gtcgacttct tcattcttctg ctaagtgttg agtatcgagt acaaattatt catcttcgtt 60  
 gtcacgttat atgagaagca tgtgtttaac attatctctt ggatattgag acggtggact 120  
 ccgataagca aacgggtcaa aagcgttttg ggcttccgta ggatactcgt acacacccgg 180  
 tgcaaccatc gtgggacttt gtggtaatga aaaaaactgt ggtggatagt tgggttgggg 240

ttgggtgttg	taaaaacccg	tttgtgga	aaatgtggtg	ggaagttggg	ttggtacaca	300
atagatccac	ctcgtctact	ttgcgagccg	cgccccctcg	cattcgaacc	ggaatccgat	360
actttttttc	ttttcatgac	ccttggtttt	tctatccatg	gtatgcatct	gaattggtaa	420
attgggtgaa	aaatggaatg	tttaacatgg	taaaatggaa	tgttataatt	ataaagggtat	480
ttaatgtttt	tttttaaacc	ataaacggtc	atatagccgt	ttaaagcaca	acgggtcaagt	540
ccccaacggg	caaatcaagc	caaacaatca	agtcgccgat	gtggcatata	tcccctgctt	600
tgaaccaggg	ccggcccaag	cctaagaatt	ttgaggcttg	ggcttttagg	ctctaaatca	660
atagggactc	taacttaaaa	aaattatata	tgatatttag	gttaagttga	atttatcttt	720
atatgtacaa	aaaaatata	ataattccaa	aataaaaaat	aaaaaaataa	attaacttcg	780
ccaacaataa	atcttttggt	atgttttgcc	ttttgttttt	ttaaaaaaag	gctcaaat	840
ttaatttgct	ttaggccacc	aaaatggtta	aaccgacctt	gctttgaacc	atacccacac	900
gacaaattag	gggatacggg	gtgggggtcg	gcaaaccaat	gccaataaac	tttgccgatc	960
aaagtttcac	catttcgggtg	attgtttgcc	gatgcggcga	ggaggaagaa	gagagaggag	1020
ggtagtgtga	ggtgggggtc	attccaatct	caactaatca	catttttttc	ctttttttta	1080
tttaaaaagt	ttaccacttc	actaaatgtc	taaaccattg	ccaacacttt	tcaccaaagt	1140
ttaaacattt	ttctctgatt	gacgtggcac	actctcattg	gttgattttt	tagtttgcca	1200
ctctcaattg	tttaaccact	ccttacaccg	tcttatgtga	gcggtggtgt	tccccaagcg	1260
acaaaagggc	ttaccgcaac	cccttaccgc	tccaacctt	tacacctta	tcttctgcct	1320
atactgcatg	tcactctatt	gcagactatc	tgagatagct	acaacctaac	gacttaaggt	1380
gaggacacgt	gtatctccaa	aaccacttcc	gtcaccctta	ccaccacgtc	atcataccac	1440
gtgccaacat	gcatgctagt	atcttctcta	ctatacacat	acttatgtac	tatatatatt	1500
caccaaatt	acatgcatgt	tagcaccctc	aattgtacaa	cacaacttaa	caacttaagt	1560
taaaagagt	aacaaaatgg	catcacaaca	gggacaacaa	acacgcaaaa	ttccggaaca	1620
ggagaagaag	gatctcgacc	aacgagcagc	taaaggcgag	accgttggtc	cgggtgttac	1680
tcgtggcaaa	tctcttgagg	ctcaagaacg	tcttgctgaa	ggtatatgtc	tatttagtta	1740
aatttacata	cgtattgggg	atggcggttt	gatatttagc	aaacggcaac	ccgaatattt	1800
aggggcccat	gaattgacat	acatatgtga	ataagattaa	ggggcctaga	ggtgtacaaa	1860
tccctttaa	aaagggtccg	gttttaaccg	gacggaaatg	gtttctaaaa	tagcgggttcg	1920
gtttttaact	ggttttacat	ggtgggcagt	cgaactggta	ctagtagatt	cggtaagcga	1980
tttgggacaa	aaaacgtttt	tttaactggt	tttgaagatc	agtaatacta	accggtatgc	2040
aaactgccag	ttcgggttga	aactggtttc	tttttctttt	tcaaaaaaga	aacggtatgg	2100
gacaaaaaac	gtttttttta	actggttttg	aagatcaata	atactaaccg	gtatgcaaac	2160
cgccagttcg	gttcgggtta	aaactggttt	tttttttttt	tttaagaag	cggtaaaaaa	2220
accggtttcg	gtcataaacc	gattttttta	acacctacaa	gggggccatg	atttaaaaaa	2280
aaagcggtaa	aaaaaaaccg	gtttcggtca	taaaccgatt	tttttaacac	ctacaagggg	2340
gccatgattt	aagatgaagt	atcggagctc	gtgatttaac	ataaagtacc	tcaaacggcc	2400
gcagtttgat	tcaagcgcaa	acccgtcggt	ctacgactag	tggcgaagct	tgagatctcc	2460
gattgggggt	cgaaaacggt	tatacccaaa	aatttctata	gaacgggggt	cgaaaacgtg	2520
tatactcaaa	aaattctata	cgaaagctac	atacctgagc	gaaaagttcg	ggtctggcac	2580
acacccctc	ccccctcttc	taacctacgc	ccctagctat	gacggtgtgg	gtagagtcca	2640
ccctttcagt	agcttttttg	cgaattcaca	ttagtttatt	ttatagttgt	agtgatgcat	2700
aataatatat	gcatgtactt	aattttgtgt	ttggtggtgg	tgaaggcg	agcaaggag	2760
gacaaacgag	gaaagaccag	ctgggaactg	aagggtacaa	ggagatggg	aaaaagggcg	2820
gtcagaccac	cggtgacaag	tcggctgggt	agcgagagg	ggaggaggag	gaggactaga	2880
tagtcaatag	tggtgtgatg	gtgtttgcat	gtacgatgat	gttaatttcc	atgttttata	2940
tatgtgtatg	tacctgtagt	atggtttagc	tcgtgtttca	tgttttgttg	gtcgttttgg	3000
tatcttcttt	agtgcagtga	cgactagtag	tcctatgatg	atgtgatgtg	atttgcatat	3060
gttgatataa	tggttaagtac	aagttttatc	tatctatctt	ttctttcggc	tagagtttac	3120
aaaatgtgta	cggatgacaa	taataagcaa	cttatatttt	tccatacaag	cttattatgt	3180
gactattgcc	caataaggca	ataagtaagt	gagccaataa	ataatagccg	gttatgtctg	3240
aacaactatt	gtgacaacaa	tattctttta	ccgaataaat	aaaaataagt	aagtattagg	3300
agccaataaa	taaataaccg	gtttaatgag	aaaagcacca	aaatagcctc	tttgatcaat	3360
agaattatga	aaatagtctc	gactcgacct	ttttgaataa	cccttcatct	cctgtgttgc	3420
gcatctctcg	cccggacgac	acgtgtctag	tgaacatgcc	tattctgatt	agtgggtgcac	3480
aaatcggggt	ttttttttaa	aaacgggttt	cgggtaggat	cgggtttttt	taaaaccggg	3540
tcgggtccgg	gtcgggtttt	ctaccaaata	tgtataccct	gtataccggg	ttcgggtcgg	3600
gttttataaa	accggtg					3617

## REIVINDICACIONES

1. Secuencia de nucleótidos constituida por el gen de girasol *Ha ds10 G1*, su promotor, intrón, secuencias 5'- y 3'-flanqueantes de *Ha ds10 G1* caracterizada porque permite la expresión génica específica de semillas.
2. Secuencia de nucleótidos según reivindicación 1, caracterizada porque está constituida por la secuencia SEQ ID N°1 y por el mapa de restricción (Figura 1).
3. Secuencia de nucleótidos caracterizada porque está constituida por fragmentos de secuencias de nucleótidos según una cualquiera de las reivindicaciones 1 a la 2.
4. Secuencia de nucleótidos, caracterizada porque comprende una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 3.
5. Secuencia de nucleótidos idéntica u homóloga a la secuencia de nucleótidos según reivindicación 4, caracterizada porque la homología es al menos en un 70%, por ejemplo en un 80% y particularmente al menos en un 95%..
6. Secuencia de nucleótidos caracterizada porque contiene secuencias según una cualquiera de las reivindicaciones 1 a la 5 en un gen quimérico.
7. Secuencia de nucleótidos según reivindicación 6, caracterizada porque permite la expresión de un gen quimérico específicamente en semillas desde etapas tempranas de maduración.
8. Secuencias de nucleótidos según reivindicación 7, caracterizadas porque están contenidas en las construcciones ds10F1, ds10F2, ds10F2Æ, ds10F3 y ds10EC1 o parte de dichas secuencias.
9. Secuencia de nucleótidos según reivindicación 7, caracterizada porque contiene secuencias codificantes y 3'-flanqueantes del gen *Ha ds10 G1*.
10. Secuencia de nucleótidos según reivindicación 9, caracterizada porque está contenida en las construcciones ds10F2 y ds10F2Æ.
11. Secuencia de nucleótidos según reivindicación 7, caracterizada porque contiene secuencias codificantes y del intrón del gen *Ha ds10 G1*.
12. Secuencia de nucleótidos según reivindicación 11, caracterizada porque está contenida en la construcción ds10F3.

13. Cassette de expresión caracterizado porque contiene una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 12.
14. Vector caracterizado porque contiene una cassette de expresión según  
5 reivindicación 13.
15. Células hospedadoras caracterizadas porque contienen una secuencia de nucleótidos según una cualquiera de las reivindicaciones 13 a la 14.
16. Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 1 a la 14 en la expresión específica de genes quiméricos en  
10 semillas, parte de semillas, en extracto de semillas, embriones de semillas y en tejidos de germínulas.
17. Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 9 a la 10 para incrementar la expresión de genes quiméricos, específicamente en semillas de plantas transgénicas.
- 15 18. Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 11 a la 12 para incrementar la expresión de genes quiméricos en semillas y/o para reducirla en otros tejidos.
19. Plantas transgénicas caracterizadas porque son plantas transformadas por una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1  
20 a la 14.



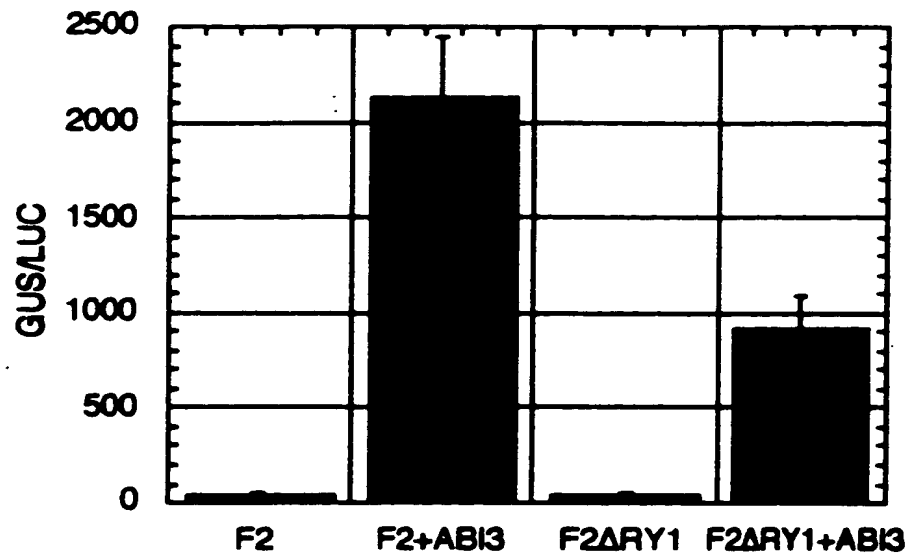


Figura 2

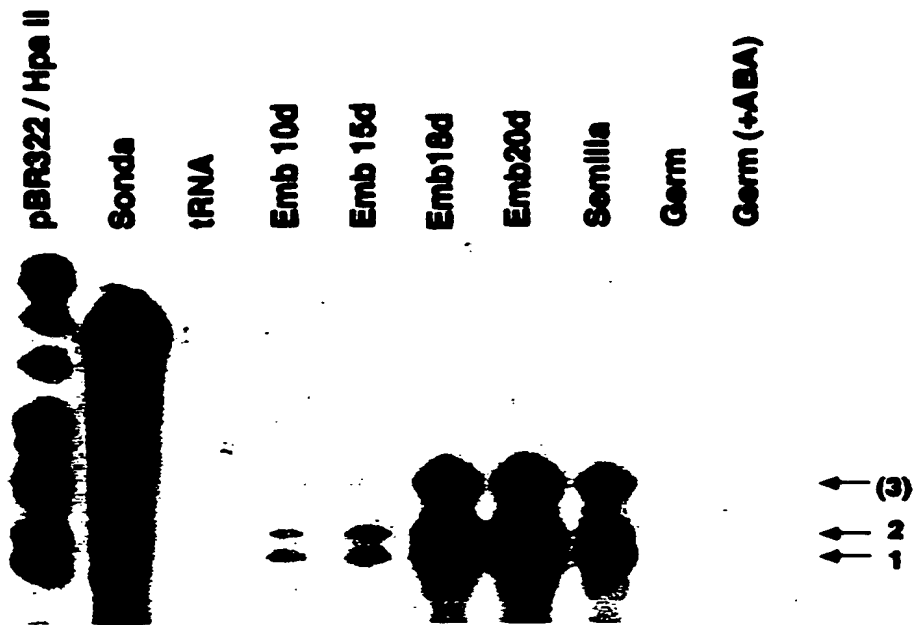


Figura 3

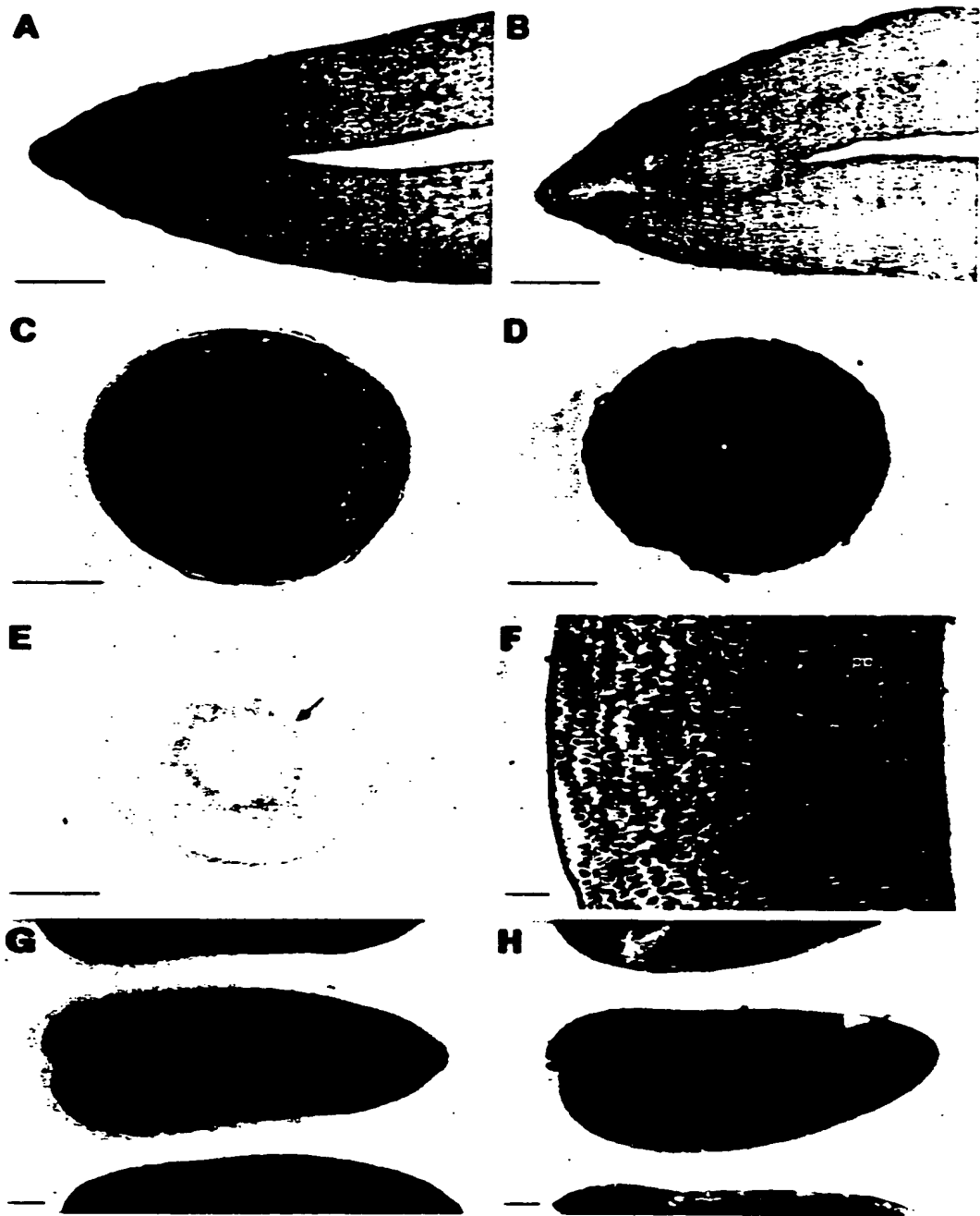


Figura 4

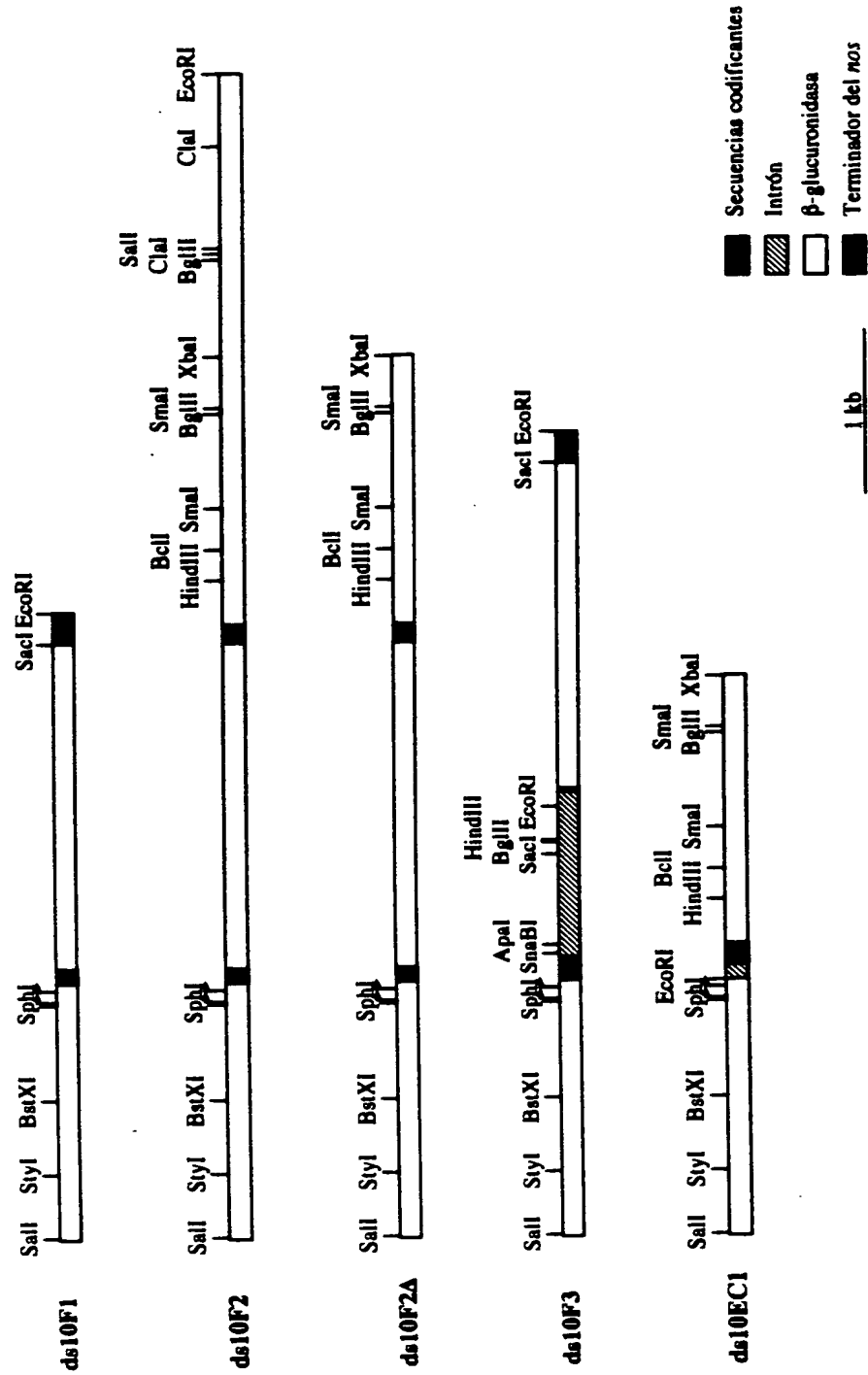


Figura 5

UNIVERSIDAD DE CALDAS

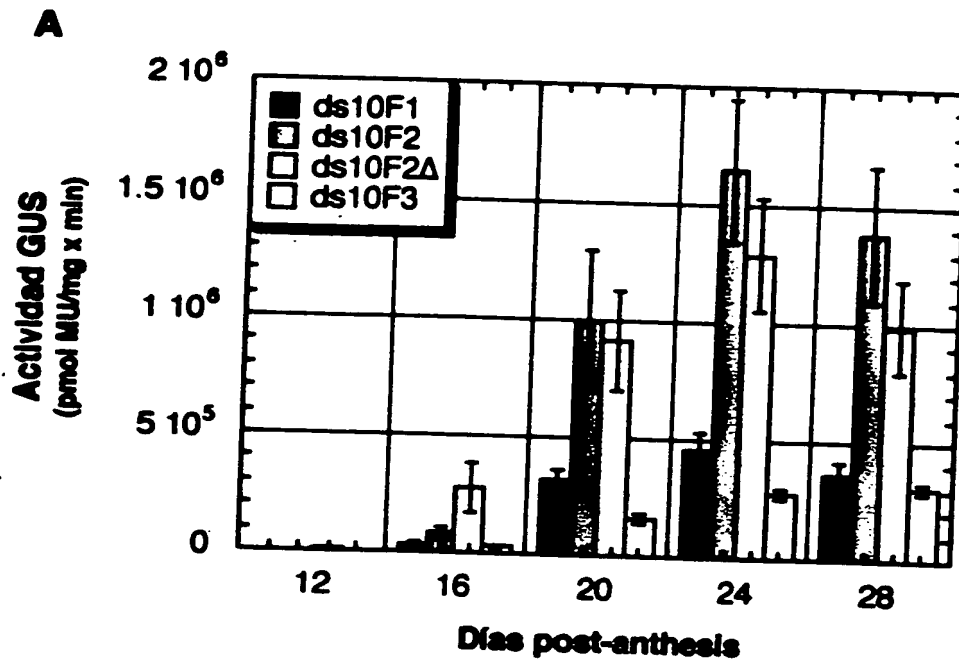
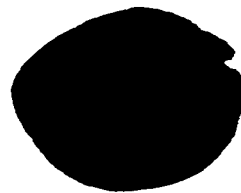
**B****C****D****E**

Figura 6